

Figure 1.

```

Sequence type explicitly set to Protein
Sequence format is Pearson
Sequence 1: new_S100_cytokine      104 aa
Sequence 2: G491246                110 aa
Sequence 3: W27152                 98 aa
Start of Pairwise alignments
Aligning...
Sequences (1:2) Aligned. Score: 23
Sequences (1:3) Aligned. Score: 34
Sequences (2:3) Aligned. Score: 29
Start of Multiple Alignment
There are 2 groups
Aligning...
Group 1:                          Delayed
Group 2:                          Delayed
Sequence:1      Score:0
Sequence:3      Score:839
Sequence:2      Score:724
Alignment Score 444
CLUSTAL-Alignment file created [/data4/genetools/lrastelli4630clustalw]

```

Multiple Alignment:

```

new_S100_cytokine  MGQCRSANAEDQEFSENERALETISIKNEHOYSWEGE - EETLEPSELRLVTQQLPHLNP
W27152            -----MAAEPSELESLETATITFFTEARQESRRKISLSVNERFELVTQQLPHLLH
G491246            -----MSOLEENIETATITFHQYSWKLESHPTLNQGEFFELVRKDLQNFLH

new_S100_cytokine  SNGC----LEEHYANLOSNDSEERSSEWE LIGEAANSVKLERP....VRSH...(SEQ ID NO:6)
W27152            DVGS----LDEHHKSLVWQDSSELKENEYWR LIGELAKEIRKED....LRIRKK.(SEQ ID NO:10)
G491246            KENKNEKVIEHIMEDLTHADKOLSEEFIMLMARLTWASHEAMHEGDEGPEHHHKPGLG

new_S100_cytokine  ----
W27152            ----
G491246            EGTP (SEQ ID NO:11)

```

Figure 2.

Multiple Alignment:

```
new_S100_cytokine      -----DNRTLKCPDTWS-ITMGQCRSANAEDAQEFSDVERAIETLIHNFHQYS
7971c.7__r0s0-212.2__EXT SISSCGAGYRTDDKTQLTEERTSWPBTMGQCRSANAEDAQEFSDVERAIETLIHNFHKYS

new_S100_cytokine      WEFGKETLTPS ELRDLVTQQLPHLMPSNCGLEEKIANLGSCNDSKLEFRSFWELIGEAAH
7971c.7__r0s0-212.2__EXT WASKKETLTPEL RDLVTQQLPHLMPSNCGLEEKIANLGNENDSKLEFCFWELIGEAAH

new_S100_cytokine      SVKMERPVGRH  (SEQ ID NO:6)
7971c.7__r0s0-212.2__EXT SVKMERPVTRS  (SEQ ID NO:3)
```

Figure 3.

hmmpfam - search a single seq against HMM database

HMMER is freely distributed under the GNU General Public License (GPL).

HMM file: pfamHMMs

Sequence file: /data4/genetools/lrastelli4423Aa315020ProteinFasta.txt

Query: AA315020

Scores for sequence family classification (score includes all domains):

Model	Description	Score	E-value	N
S_100	S-100/ICaBP type calcium binding domain	40.9	2.9e-08	1

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
S_100	1/1	32	74 ..	1	44 []	40.9	2.9e-08

Alignments of top-scoring domains:

S_100: domain 1 of 1, from 32 to 74: score 40.9, E = 2.9e-08

```

      *->LEKkaietiInvFhqYSgreGdkdtLsKkELKellekELpnflkn<-*
          K+aiet+I+ FhqYS eG k tL+ EL+ L++++Lp+ +
AA315020  32  VERAIETLIKRFHQYS-VEGCKETLTPSELRLDVTQQLPHLMPS  74(SEQ ID NO:33)

```

//

Figure 4B

Table 3
AY007220
Consensus

10	20	30	40	50	60
.....					
MGQCRSANAEDAQEFSDVERAIETLIKNFHRYSVASKKETLTPELRDLVTQOLPHLMPS					
MGQCRSANAEDAQEFSDVERAIETLIKNFHRYSVGCKETLTPELRDLVTQOLPHLMPS					
MGQCRSANAEDAQEFSDVERAIETLIKNFHYSVSKETLTPELRDLVTQOLPHLMPS					

Table 3
AY007220
Consensus

70	80	90	100
.....			
NCGLEEKIANLGNCNDSKLEFCSFWELIGEAAKSVKERPV (SEQ ID NO:3)			
NCGLEEKIANLGSCNDSKLEFRSFWELIGEAAKSVKERPV (SEQ ID NO:39)			
NCGLEEKIANLGNCNDSKLEFSFWELIGEAAKSVKERPV (SEQ ID NO:40)			

Figure 4A.

BLOCKS Protein Domain Analysis of new_S100_cytokine

Probe Size: 104 Amino Acids

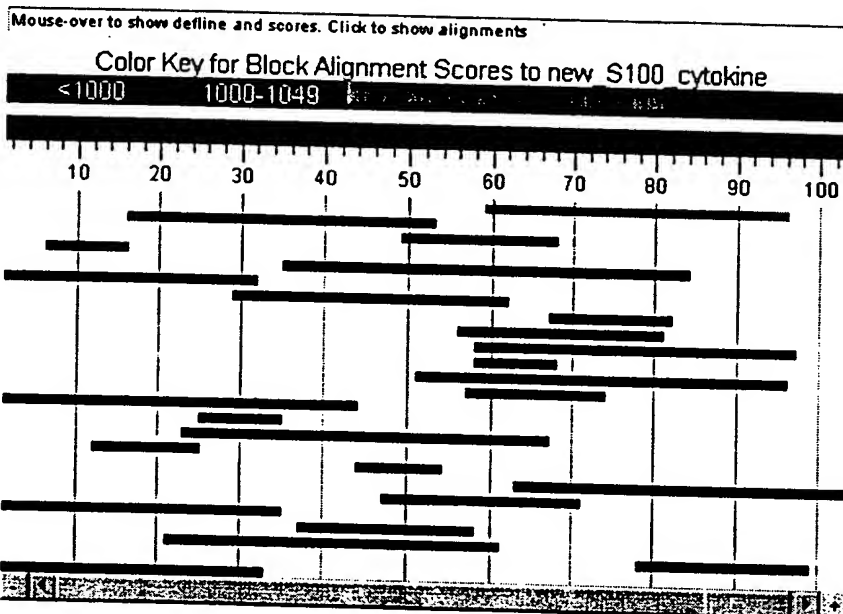
Probe File: lrastelliblocks.seq

Target File (s) : blocks.dat

Records Searched: 4034

Scores Done: 4034

Alignments Done: 535470



Accession	Description	Strength	Score	RT	Accession	Description
EL00202D	3-100/100 type calcium binding protein.	1236	1057	0	59	NCLEXIAHLGSDNHLFSTFGLICAAK (SEQ ID NO:12)
EL00202A	3-100/100 type calcium binding protein.	1245	1020	0	18	D-ERAILTLIK-FTSTFVEGLKLTIPKLRDVEQ (SEQ ID NO:13)
EL00074A	Bacterial type II secretion system protein F	1456	1021	0	49	VTQQLPMLPINCGLKEDI (SEQ ID NO:14)
EL00072B	Ubiquitin carboxyl-terminal hydrolases family	1217	991	0	6	AAALDAQET (SEQ ID NO:15)
EL00538C	Bacterial chemotaxis sensory transducers prot	1750	990	0	25	GGGDLTPSLEEDIVTQQLPMLPINCGLKEDI (SEQ ID NO:16)
EL00532I	Phosphoenolpyruvate carboxykinase (ATP) prote	1412	989	0	0	mgqELHnde dAQe13GTERAIDLIDNThq (SEQ ID NO:17)
EL00704A	Prokaryotic-type carbonic anhydrases proteins	1539	987	0	29	hqt3vqgKrcLcPseLr dLVTQQLPMLPINCGLKEDI (SEQ ID NO:18)
EL01017E	Ergosterol biosynthesis ERG4/ERG14 family pro	1499	983	0	67	lanlqLCHdKLETR (SEQ ID NO:19)
EL00210E	Lysosome-associated membrane glycoproteins du	1693	976	0	56	LPPLKGLKEDIHLGSDNHLFSTFGLICAAK (SEQ ID NO:20)
EL00422C	Phosphotransferase proteins.	1501	976	0	58	plnLGLKEDIHLGSDNHLFSTFGLICAAK (SEQ ID NO:21)
EL00002	PH domain proteins profile.	990	976	0	58	PERGLKEDI (SEQ ID NO:22)
EL00459	Myotomus proteins.	1273	974	0	61	qqlpLMLPINCGLKEDIHLGSDNHLFSTFGLICAAK (SEQ ID NO:23)
EL00007C	Phosphatidylinositol-specific phospholipase X	1493	974	0	57	mPlnLGLKEDIHLGSDNHLFSTFGLICAAK (SEQ ID NO:24)
EL01072B	Glypican proteins.	1500	973	0	12	qLIDVRAITL (SEQ ID NO:25)
EL00279B	Membrane attack complex components / perforin	1117	972	0	25	AKNThqVE (SEQ ID NO:26)
EL01110D	Urease nickel ligands proteins.	1592	970	0	23	TLIDNThqVEgKLTIPKLRDVEQQLPMLPINCGLKEDI (SEQ ID NO:27)
EL00175D	Phosphoglycerate mutase family phosphohistidi	1290	960	0	12	qLIDVRAITL (SEQ ID NO:28)
EL00579A	Ribosomal protein L19 proteins.	1092	967	0	44	ELNHLVQGL (SEQ ID NO:29)
EL00022D	2'-5'-oligoadenylate synthetases proteins.	1026	967	0	63	LeckiaHLGSDNHLFSTFGLICAAK (SEQ ID NO:30)
EL01005C	Formate and nitrite transporters proteins.	1562	964	0	67	GLVTqLPLMLPINCGLKEDIHLGSDNHLFSTFGLICAAK (SEQ ID NO:31)
EL00151A	Glycoprotein hormones beta chain proteins.	1520	963	0	1	qQCRdHnde dAQETDVEALITLIDNThqVE (SEQ ID NO:32)
EL00653A	Vinculin family calin-binding region proteins	1567	963	0	37	hLTPSLr dLVTQQLPMLPINCGLKEDIHLGSDNHLFSTFGLICAAK (SEQ ID NO:33)
EL00745A	Beta-lactamases class B proteins.	1580	960	0	21	LeTLIDNThqVEgKLTIPKLRDVEQQLPMLPINCGLKEDIHLGSDNHLFSTFGLICAAK (SEQ ID NO:34)
EL01021B	Heat shock hsp10 proteins family profile.	1200	960	0	70	LEFHTFGLICAAK (SEQ ID NO:35)
EL01062C	Hydroxymethylglutaryl-coenzyme A lyase protei	1900	960	0	-1	HqqrLr dHnde dAQETDVEALITLIDNThqVE (SEQ ID NO:36)

Figure 4B.

Table 3
AA007220
Consensus

```

      10      20      30      40      50      60
.....|.....|.....|.....|.....|.....|
MGQCRSANAEDAQEFSDVERAIETLIKNNFHYSVAGKKTLPSELRLDLVTQQLPHLMPS
MGQCRSANAEDAQEFSDVERAIETLIKNNFHYSVAGKKTLPSELRLDLVTQQLPHLMPS
MGQCRSANAEDAQEFSDVERAIETLIKNNFHYSVAGKKTLPSELRLDLVTQQLPHLMPS

```

Table 3
AA007220
Consensus

```

      70      80      90      100
.....|.....|.....|.....|
NCGLEEKIANLGNNDKSKLEFSFWELIGEAAKSVKLERPV (SEQ ID NO:3)
NCGLEEKIANLGNNDKSKLEFSFWELIGEAAKSVKLERPV (SEQ ID NO:39)
NCGLEEKIANLGNNDKSKLEFSFWELIGEAAKSVKLERPV (SEQ ID NO:40)

```

Figure 4C.

Table 6
AA007220
Consensus

```

      10      20      30      40      50      60
.....|.....|.....|.....|.....|.....|
MGQCRSANAEDAQEFSDVERAIETLIKNNFHOYSVEGGKETLTPSELRLDLVTQQLPHLMPS
MGQCRSANAEDAQEFSDVERAIETLIKNNFHOYSVEGGKETLTPSELRLDLVTQQLPHLMPS
MGQCRSANAEDAQEFSDVERAIETLIKNNFHOYSVEGGKETLTPSELRLDLVTQQLPHLMPS

```

Table 6
AA007220
Consensus

```

      70      80      90      100
.....|.....|.....|.....|
NCGLEEKIANLGNNDKSKLEFSFWELIGEAAKSVKLERPVRGH (SEQ ID NO:6)
NCGLEEKIANLGNNDKSKLEFSFWELIGEAAKSVKLERPVRGH (SEQ ID NO:39)
NCGLEEKIANLGNNDKSKLEFSFWELIGEAAKSVKLERPVRGH (SEQ ID NO:40)

```

Figure 4D.

Table 3
gi|4139958|pdb|1MHO|
PROTEIN MRP-126
ICTACALCIN
CALGRANULIN B
Consensus

```

      10      20      30      40
.....|.....|.....|.....|
EKAIVVALIDVPHOYSREGDKHKLKSELRLDLVTQQLPHLM (SEQ ID NO:3)
EKAIVVALIDVPHOYSREGDKHKLKSELRLDLVTQQLPHLM (SEQ ID NO:41)
EKAIVVALIDVPHOYSREGDKHKLKSELRLDLVTQQLPHLM (SEQ ID NO:42)
EKAIVVALIDVPHOYSREGDKHKLKSELRLDLVTQQLPHLM (SEQ ID NO:43)
EKAIVVALIDVPHOYSREGDKHKLKSELRLDLVTQQLPHLM (SEQ ID NO:44)
EKAIVVALIDVPHOYSREGDKHKLKSELRLDLVTQQLPHLM (SEQ ID NO:45)

```

Figure 4E.

Table 6
gi|4139958|pdb|1MHO|
PROTEIN MRP-126
CALGRANULIN B
CALGRANULIN B
Consensus

```

      10      20      30      40
.....|.....|.....|.....|
EKAIVVALIDVPHOYSREGDKHKLKSELRLDLVTQQLPHLM (SEQ ID NO:6)
EKAIVVALIDVPHOYSREGDKHKLKSELRLDLVTQQLPHLM (SEQ ID NO:41)
EKAIVVALIDVPHOYSREGDKHKLKSELRLDLVTQQLPHLM (SEQ ID NO:42)
EKAIVVALIDVPHOYSREGDKHKLKSELRLDLVTQQLPHLM (SEQ ID NO:44)
EKAIVVALIDVPHOYSREGDKHKLKSELRLDLVTQQLPHLM (SEQ ID NO:46)
EKAIVVALIDVPHOYSREGDKHKLKSELRLDLVTQQLPHLM (SEQ ID NO:47)

```

Figure 5

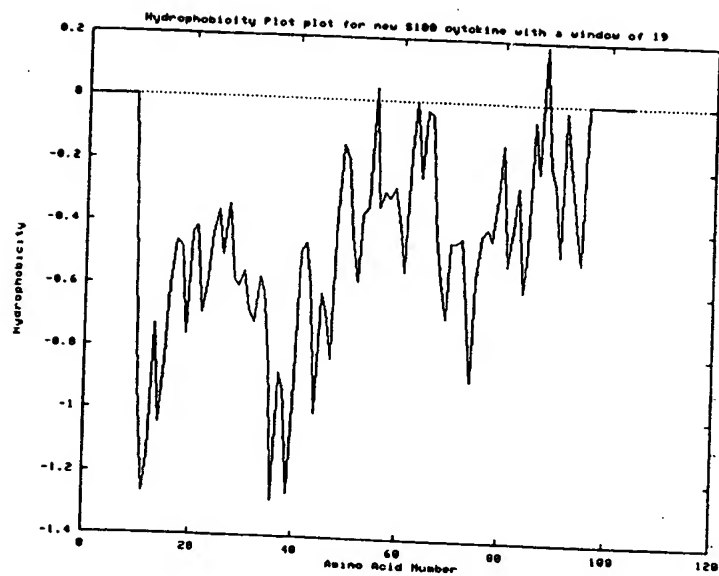


Figure 6

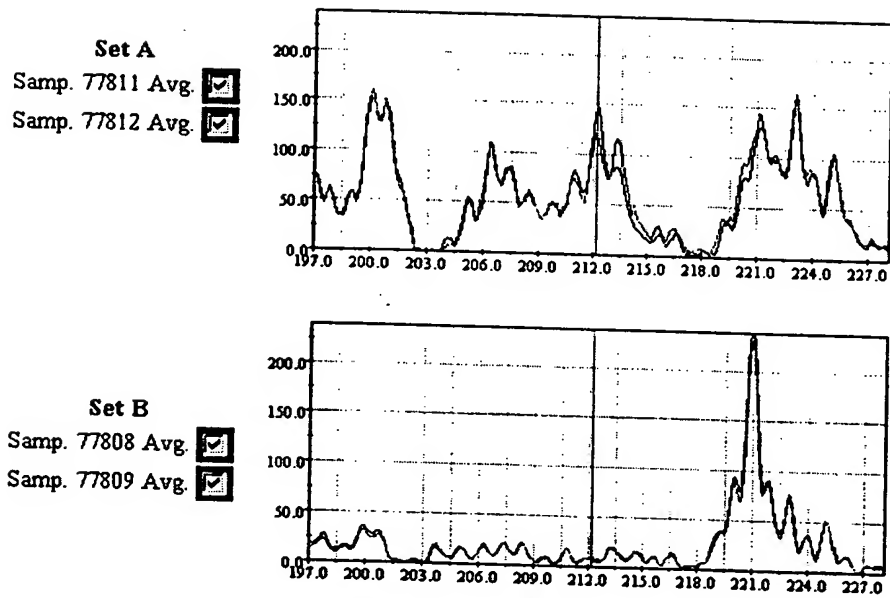


Figure 7

***** Contig 1 *****	
65677221+	GAATTCCAGAGGGAGTTCTCAGTGCCCCGGACAGGCCTCTCCAGCTTCACACTCTTGGC
AA315020-	TGCCCCCGGACAGTCCTCTCNAGCTTCACACTCTTGGC
consensus	GAATTCCAGAGGGAGTTCTCAGTGCCCCGGACAGGCCTCTCCAGCTTCACACTCTTGGC
65677221+	CGCTTCTCCAATCAGCTCCCAGAAACTCCTGAACTCCAGTTTAGAGTCATTGCAGCTGCC
AA315020-	CGCTTCTCCAATCAGCTCCCAGAAACTCCTGAACTCCAGTTTAGAGTCATTGCAGCTGCC
consensus	CGCTTCTCCAATCAGCTCCCAGAAACTCCTGAACTCCAGTTTAGAGTCATTGCAGCTGCC
65677221+	CAGGTTGGCAATTTTCTCTTCCAGGCCACAGTTGCTCGGCATGAGATGGGGCAGCTGCTG
AA315020-	CAGGTTGGCAATTTTCTCTTCCAGGCCANAGTTGCTCGGCATGAGATGGGGCAGCTGCTG
consensus	CAGGTTGGCAATTTTCTCTTCCAGGCCACAGTTGCTCGGCATGAGATGGGGCAGCTGCTG
65677221+	GGTGACCAGGTCCCGTAGCTCAGAAGGGGTCAGCGTCTCCTTCCCACCCTCCACGGAGTA
AA315020-	GGTGACCAGGTCCCGTAGCTCAGAAGGGGTCAGCGTCTCCTTCCCACCCTCCACGGAGTA
consensus	GGTGACCAGGTCCCGTAGCTCAGAAGGGGTCAGCGTCTCCTTCCCACCCTCCACGGAGTA
65677221+	CTGGTGAAAGTTCTTGATGAGGGTCTCAATGGCCCTCTCCACATCACTGAATTC (SEQ ID NO:37)
AA315020-	CTGGTGAAAGTTCTTGATGAGGGTCTCAATGGCCCTCTCCACATCACTGAATTCCTGAGC
consensus	CTGGTGAAAGTTCTTGATGAGGGTCTCAATGGCCCTCTCCACATCACTGAATTCCTGAGC
AA315020-	ATCCTCTGCGTTGGCTGACCGACACTGTCCCATGGTGCTCACTGTGTCTGGTCCTTTGGT
consensus	ATCCTCTGCGTTGGCTGACCGACACTGTCCCATGGTGCTCACTGTGTCTGGTCCTTTGGT
AA315020-	GAGAGTTCTGTTGTCCTAT (SEQ ID NO:4)
consensus	GAGAGTTCTGTTGTCCTAT (SEQ ID NO:5)